JOURNAL OF BACTERIOLOGY, Apr. 1994, p 2265–2271
0021-9193/94/$04.00+0

Vol. 176, No 8

Tn*10*-Mediated Inversions Fuse Uridine Phosphorylase (*udp*) and rRNA Genes of *Escherichia coli*

MICHAEL FONSTEIN,†* TATIANA NIKOLSKAYA,† DIMITRIY ZAPOROJETS, YURI NIKOLSKY,† SAULIUS KULAKAUSKAS,‡ AND ALEXANDER MIRONOV

Institute of Genetics and Selection of Industrial Microorganisms, Moscow, Russia

Received 12 October 1993/Accepted 10 February 1994

Two strains carrying *metE*::Tn*10* insertions (upstream of the *udp* gene) were used to isolate mutants of *Escherichia coli* overexpressing *udp*. These strains differ in their gene order; one contains an inversion between the *rrnD* and *rrnE* rRNA operons. Selection was based on the ability of overexpressed Udp to complement thymine auxotrophy. Chromosomal rearrangements that connect the *udp* gene and promoters of different *rrn* operons were obtained by this selection. Seven of 14 independent mutants selected in one of the initial strains contained similar inversions of the *metE-rrnD* segment of the chromosome (about 12% of its length). Another mutant contained traces of a more complicated event, inversion between *rrnB* and *rrnG* operons, which was followed by reinversion of the segment between *metE* and the hybrid *rrnG/B* operon. Similar inversions (*udp-rrn*) in a strain already carrying an *rrnE-rrnD* inversion flip the chromosomal segment between *metE* and *rrnD/E* in the opposite direction. In this case, inversions are also accompanied by duplications of the chromosomal region between the *rrnA* and hybrid *udp-rrnD/E* operons. PCR amplification with a set of oligonucleotides from the *rrn*, Tn*5*, and *met* genes was used for more detailed mapping. Amplified fragments of the rearranged chromosomes connecting *rrnD* sequences and insertion elements were sequenced, and inversion endpoints were established.

The genomes of *Escherichia coli* and *Salmonella typhimurium* have an extremely conservative arrangement of genes (25). The only striking difference between them is a large inversion of approximately 15% of the chromosome (26), although inversions are considered to be rarer than other rearrangements of the bacterial genome. For example, no inversions were detected among 1,500 random *his* auxotrophs, which were analyzed by Hartman et al. (10). It was not possible to find a chromosomal inversion between *lac* and φ80 *att* by using the powerful selection for the restoration of the Lac$^+$ phenotype (16). The majority of artificially constructed *E. coli* strains with extended chromosomal inversions show low viability (11); however, naturally isolated inversion strains, such as *E. coli* W2637 and its derivative W3110, are rather stable (12). One possible explanation of these facts is that the actual copy numbers of different chromosomal segments can measurably differ, regardless of their position in the bacterial chromosome. Inversions may vary the expression of genes by changing their distances (and copy numbers) from the chromosomal Ori. Most of the stable inversions are symmetrical to the chromosomal Ori (or Ter) and have less chance to affect gene expression and inhibit cell growth (13). Stabilization of the genome in strains carrying artificial inversions occurs by reinversions which restore the original gene order. These reinversions may produce translocations (11). Areas of the *E. coli* genome in which the ends of inversions are prohibited were localized by Rebollo et al. (24, 29) and are summarized elsewhere (19).

Although they are rare in general, certain inversions are frequent enough to make a significant contribution to *E. coli* genome variation. For example, inversions between *rrn* operons occur at about 10^{-5} (12). DNA rearrangements induced by *rrn* DNA repeats are revealed in various microorganisms (for example, see reference 26). These inversions can play a significant role in the evolution of the genome, first destabilizing it and then inducing a sequence of stabilizing events.

We have described a selection procedure for the isolation of *E. coli* variants with a high level of uridine phosphorylase (*udp*) gene expression. The genetic analysis of one selected strain (CM973) revealed that it carried a large inversion in which one breakpoint was located within the *metE* gene, while the other breakpoint was mapped between 72 and 73 min on the genetic map (17). On the basis of both genetic and physiological behaviors of the inversion mutant CM973, we suggested that the non-*metE* endpoint of the inversion was located within the rRNA operon *rrnD* and that a high level of *udp* gene expression in this mutant was due to read-through transcription from the promoters of the *rrnD* operon.

To confirm the results of the genetic mapping and for further study of other selected variants, the chromosomal DNAs of the *E. coli* strains isolated by the same selection procedure from a different genetic background were physically mapped. The construction of a *Not*I physical map of *E. coli* (31) allows the genetic mapping of this organism to be done directly by observing alterations in the mobilities of DNA fragments on pulsed-field gels. This low resolution mapping can be continued by using PCR (with subsequent sequencing of its products) to reveal exact nucleotides affected by the rearrangements studied. The data achieved by this approach show that *rrn* operons are involved in generation of the Tn*10*-promoted rearrangements found in all of the inversion type mutants selected for Udp overexpression.

MATERIALS AND METHODS

Strains, plasmids, and culture conditions. The bacterial strains and plasmids used in this study and their genetic

* Corresponding author. Phone: (312) 702-1081. Fax: (312) 702-3172. Electronic mail address: m-fonstein@uchicago.edu.

† Present address: Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58 Street, Chicago, IL 60637.

‡ Present address: Institute National de la Recherche Agronomique, Domaine de Vilvert, 78352 Jouy-en Josas, France.

TABLE 1. *E. coli* K-12 strains used in the work

Strain	Genotype	Comment
CM922	Hfr3.0S0 *thi thyA deoC deoA metE*::Tn*10*	
AM174	CM922 *metE*$^+$	
CM973	CM922 *INV1* [*metE' rrnD'*]	
AM2217	CM973 *argG*::Tn*5*	
AM2237	CM973 *cysG*::Tn*5*	
AM2222	CM973 *metE*::Tn*5*	Used to generate data in Fig. 1 and not mentioned in the text
AM2224	CM973 *aroE rpsL*	
CM991	CM973 *zif-9*::Tn*10*	Used to generate data in Fig 1 and not mentioned in the text
CM1051	CM973 *udp*::Tn*5*	Used to generate data in Fig 1 and not mentioned in the text
AM2302	CM1051 *zif-9*::Tn*10*	Used to generate data in Fig. 1 and not mentioned in the text
AM2301	CM973 *rpsE*	Used to generate data in Fig 1 and not mentioned in the text
AM2113	CM922 *INV*157 [*metE' rrnD'*]	
AM2118	CM922 *INV*158 [*metE' rrnD'*]	
AM2129	CM922 *INV*159 [*metE' rrnD'*]	
AM2131	CM922 *INV*160 [*metE' rrnG/B*]	
AM2138	CM922 *INV*161 [*metE' rrnD'*]	
AM2141	CM922 *INV*162 [*metE' rrnD'*]	
AM2144	CM922 *INV*163 [*metE' rrnD'*]	
AM2150	W3110 *thi thyA deoC deoA metE*::Tn*10*	
AM2151	AM2150 *INV/DUP30*	
AM2152	AM2150 *INV/DUP31*	
AM2153	AM2150 *INV/DUP33*	
AM2154	AM2150 *INV/DUP34*	
AB1157	F$^-$ *thr-1 ara-14 leuB6 (gpt-proA)62 sx-33 supE44 galK2 hisG4 rfbD1 mgl-51 rpsL31 kdgK51 xyl-5 mtl-1 argE3 thi-1*	

characteristics are presented in Table 1. Growth conditions for *E. coli* strains and the concentrations of growth factors and antibiotics have been described elsewhere (21, 22). Phage crosses were done according to Miller (21) for Mu and according to Young et al. (33) for T4 and are described in detail by Mironov and Sukhodolets (22).

DNA preparation. Alkaline plasmid extractions for labeling were done according to the method described by Maniatis et al. (20). Genomic DNA of *E. coli* was prepared by sodium dodecyl sulfate lysis in the presence of proteinase K, and then by phenol extraction. Samples of chromosomal DNA for pulsed-field gel electrophoresis (PFGE) were prepared by embedding bacterial cells in 0.6% agarose blocks at a final cell concentration of 5×10^9 cells per ml, which is equivalent to 1 mg of DNA per 10 ml of agarose matrices. Lysis was achieved by 1% lauryl sarcosine–1 mg of proteinase K per ml in 0.2 M EDTA (pH 8.0) at 55°C for 14 to 36 h (9).

Enzymatic manipulation of DNA. Restriction endonucleases and the Klenow fragment of DNA polymerase were purchased from Fermentas (Vilnius, Lithuania) and were used according to the instructions of the manufacturer. For the cleavage of embedded samples of DNA used in PFGE, 5 to 20 U of *Not*I and 3-h incubations at 37°C were used.

Electrophoresis. PFGE was performed with the Diagnosticum system. Standard running conditions were as follows: 24 h, 0.5× Tris-borate-EDTA (19), 1% BRL agarose, 8 to 14°C running temperature, and 10-V/cm field strength. To resolve restriction fragments between 40 and 400 kb, switching intervals of 25 s were used. To resolve 500- to 1,000-kb fragments, the switching intervals were 100 s (9). After electrophoresis, the gels were stained with ethidium bromide and photographed.

Oligonucleotides, PCR amplification, and sequencing. The following oligonucleotides were used for PCR and sequencing: rrn1 (−82 to −61) (*rrn* coordinates are relative to the P1 transcription start of *rrnB* operon [5]), CTGATTTGGTTG AATGTTGCGC; rrn2 (733 to 712), AAAGTACTTTA CAACCCGAAGG; rrn3 (741 to 762), AGGAAGGGAGTAA AGTTAATAC; rrn4 (1737 to 1716), GCTACCTACTCTTT TGCAACC; rrn5 (1753 to 1774), CGCTTACCACTTGT GATTCAT; rrn6 (3317 to 3294), CCTTCCCACATCGTTTC CCACTTA; rrn7 (3338 to 3351), GGCTTAGAAGCAGC CATCATTT; rrn8 (4378 to 4356), CTCAATGTTCAGTGT CAAGCTAT; rrn10 (5671 to 5650), CGGTTCATTAGTAC CGGTTAGC; IS_p (external end of the IS*10* with coordinates 152 to 128), GCAGAATTGGTAAAGAGAGTCGTG; IS_d (internal end of the IS*10* with coordinates 1178 to 1203), GTACTCTCAACAGTTCGCTTAGGCA; $metE_p$ (promoter-proximal region of the *metE* gene), GGCGGTAGGGCGT GAATTGCG; $metE_d$ (distal region of the *metE* gene), CCC GACGCAAGTTCTGCGCCGC.

PCR amplification consisted of 30 cycles of the following steps: 95°C for 30 s, 55°C for 40 s, and 72°C for 90 s. Other conditions were as recommended by Perkin-Elmer Cetus, producer of the *Taq* polymerase that was used.

The products of PCR, which link *rrn* DNA to the IS*10* end, allow sequencing of both strands with IS_d and the nearest *rrn* primers. DNA fragments were purified by the Bio 101 Geneclean kit and were used with ^{32}P-end-labeled primers in a linear amplification sequencing protocol (23).

Blot hybridization. DNA fragments from pulsed-field gels were transferred onto GeneScreen Plus nylon membranes by the standard capillary procedure (20) and hybridized with 0.5×10^6 to 2×10^6 cpm of a randomly labeled probe per ml. Hybridization, washing, and removal of the probe were done according to the membrane manufacturer's protocol. The only

exception was the final 0.1 × SSC (1 × SSC is 0.15 M NaCl plus 0.015 M sodium citrate) wash, which was done at 65°C. Filters were exposed to Kodak X-ray film from 2 to 48 h. By being kept wet, each filter could be used for five hybridizations, with the previous probe removed each time.

RESULTS

Selection procedure for the isolation of inversion mutants. The *udp* gene, located at 85 min on the *E. coli* genetic map, encodes a protein catalyzing the phosphorolysis of uridine and also has some specificity for thymidine. Udp is able to convert the thymine to thymidine complementing a *thyA* mutation under the conditions of its constitutive synthesis in a strain deficient in thymidine phosphorylase (*deoA*). Promoter operator constitutive mutations or duplications of the *udp* gene were isolated in this way (1, 22).

udp gene expression also depends on the transcription of an adjacent *metE* gene (1). Because of this, enhancement of *udp* expression may be achieved by deleting the putative *rho*-dependent terminator between the *udp* and *metE* genes. To study such a variant, we constructed the strain CM922 containing *thy deoA* mutations and a Tn*10* insertion in the distal part of the *metE* gene. Fourteen independent thymine-utilizing mutants showing sensitivity to tetracycline and retaining methionine auxotrophy were isolated. Surprisingly, 8 of 14 Tet^s *metE* mutants could not be transduced to MetE$^+$ in P1 transduction crosses with *metE*$^+$ donor strains. The following genetic analysis of one of these mutants (CM973) indicated that it contained an extended inversion, supposedly providing new strong promoters for *udp* expression. This strain and seven others (named AM2113, AM2118, AM2129, AM2131, AM2138, AM2141, and AM2144) were further analyzed by means of physical mapping.

A second parental strain (AM2150) was used for isolation of mutants by the same selection procedure. This strain, a derivative of W3110, already contained an inversion between the *rrnD* and *rrnE* operons and therefore had an orientation opposite to that of the *udp* gene relative to the *rrn* operons. Four independent mutants (AM2151 to AM2154) were selected on the basis of an increased level of *udp* gene expression.

Genetic analysis of the *INV1* inversion mutant. Genetic analysis of strain CM973, which carried the inversion designated *INV1*, demonstrated that the *udp* gene was linked in P1 transduction crosses with the markers *crp* and *rpsE* located at 73 min on the *E. coli* map (2). We used the marker *aroE* located between *rpsE* and *rrnD* for more precise mapping of *INV1* (Fig. 1).

To determine the breakpoint locations of *INV1* on the chromosome, four-factor transduction crosses were performed by using P1 phage propagated on the donor strain AM2302 containing a *zif9*::Tn*10* insertion near *udp* (the results of crosses are presented in Fig. 1). The Tet^r marker was transduced into the *aroE udp*$^+$ Tet^s recipient (strain AM2224 = CM973 [*aroE rpsL*]). A total of 225 Tet^r transductants were examined for the inheritance of *udp*, *aroE*, and *rpsE* (Sp^r) as unselected markers. These results (Fig. 1) indicated that *aroE* is located between the *rpsE* and *udp* markers, showing 90% linkage with *udp*. This means that the non-*metE* endpoint of the *INV1* inversion lies approximately in the area of the *rrnD* operon (Fig. 1).

The other breakpoint of this inversion, marked by a *metE*::Tn*5* insertion, was linked to the *argG* marker located at 69 min on the *E. coli* map (2% linkage) (detailed description of the crosses in reference 17). This was done in transduction crosses with phage T4 (GT7) (33); this phage was chosen for its

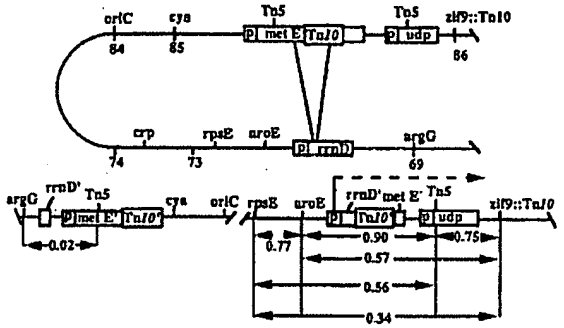

FIG. 1. Proposed scheme of the *INV1* inversion. (A) Gene map of the parental strain CM922 and a scheme of the intramolecular Tn*10*-mediated inverse transposition; (B) gene map of the rearranged variant CM973. Numbers under the map in panel A represent gene positions according to the standard genetic map of *E. coli*. Numbers inside the lines with arrows in panel B represent genetic linkage (cotransductional frequency) measured in P1 transduction experiments. Genes inside the bars are the substrate of the rearrangement studied.

transduction capacity, which is higher than that for P1. The weak coinheritance (2%) observed may also be simulated by secondary transposition (zygotic induction during DNA transfer) in the absence of any real linkage. However, further physical mapping and sequencing were consistent with the proposed *argG-rrnD'-metE'* linkage (Fig. 1).

These results, together with the data showing the growth rate dependence of *udp* expression in the CM973 mutant (17), allow us to propose that the non-*metE* breakpoint of the *INV1* inversion is located close to the *rrnD* operon, connecting the *rrnD* promoters with the *udp* gene.

Alignment of *Not*I restriction pattern of CM922 (HfrH). Since the parental strain CM922 (AM174 *metE*, a derivative of HfrH) used in this work may contain unrecognized chromosomal deviations from the *E. coli* strains that were already physically mapped, it was necessary to compare its *Not*I pattern with those of AB1157, AM2150, and W3310, which had been mapped already (7, 31, 32). A comparison of the restriction patterns obtained from these strains revealed some differences (data not shown) due to the integration of the F factor around 97 min of the standard genetic map in the strain CM922 (18) and to the presence of an extrachromosomal copy of the F plasmid. Another difference between the strains is that fragment I (*Not*I fragments numbered according to Smith et al. [31]) is larger and fragment D is smaller in strain AB1157 than in strain AM174 (a direct ancestor of CM922). Differences in sizes for the other *Not*I fragments do not exceed 3% (limits of resolution). This pattern comparison confirmed the possibility of applying the *Not*I restriction analysis to the strains AM174 and CM922 to map their rearrangements.

Physical mapping of the inversion mutants derived from CM922. Chromosomal DNAs from the strains selected from CM922, namely, CM973, AM2113, AM2118, AM2129, AM2131, AM2138, AM2141, and AM2144, were treated with *Not*I and separated by PFGE. The restriction patterns of all clones except AM2131 were identical. For more precise mapping, we constructed two additional derivatives of strain CM973 containing the insertions *cysG*::Tn*5* (75 min) and *argG*::Tn*5* (69 min), strains AM2237 and AM2217, respectively. The *Not*I sites introduced into a known map position as a part of the insertion elements of Tn*5* (32) split the large *Not*I

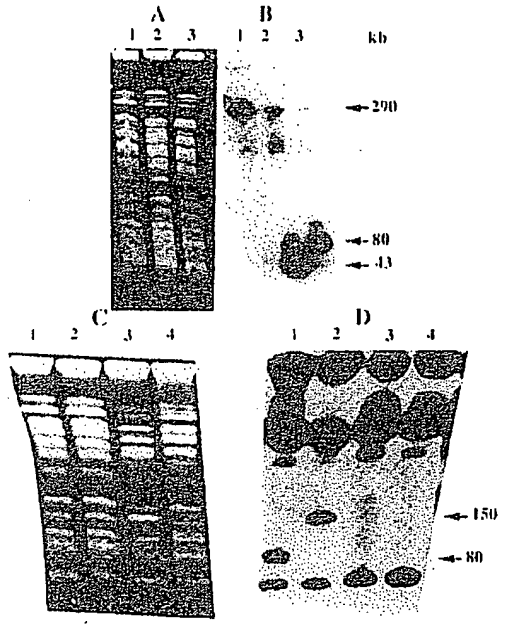

FIG 2 Localization of the *udp* gene on the *Not*I restriction fragments of the different *E. coli* strains (A and B) lanes 1, CM973 (*INV*); lanes 2, CM922 (parent); lanes 3, AM2237 (*INV cysG*::Tn*5*); (C) lane 1, CM922; lane 2, CM973; lane 3, AM2217 (*INV argG*::Tn*5*); lane 4, AM2237; (D) lane 1, AM2237; lane 2, AM2217; lane 3, CM973; lane 4, CM922. PFGE separation (A and C) and blot hybridization with the *udp* probe (B) and the *rrn* probe (D) are shown.

fragments, making the size determination more accurate. The plasmids pUD7 (*udp* [4]) and pKK3535 (*rrnB* [14]) were used as probes to determine the locations of the *udp* gene and the *rrn* operons on the macrofragments (Fig. 2). The observed changes in mobilities of the *Not*I fragments harboring the *udp* and *rrn* genes revealed on the gel blots generated a map that is similar for seven strains (Fig. 3). The data confirmed that one of the proposed breakpoints of *INV*1 and the other six inversion mutations are within fragment A (in or near *rrnD*), while the opposite breakpoint is within the small fragment S (in or near *metE*) (Fig. 3). Such an inversion would form two new *Not*I fragments: A′ (about 750 kb) and S′ (290 kb) instead of A (1,000 kb) and S (43 kb) of the parental strain CM922 (AM174). This was demonstrated by PFGE analyses and blot hybridizations.

Size determination errors that were reduced to 5 to 10 kb by application of the AM2237 and AM2217 strains made it possible to map rearrangement endpoints in the vicinity of the *rrn* and *udp* genes with the accuracy of a few kilobases. More precise mapping (revealing endpoints within *rrn* genes with nucleotide resolution) was done by PCR mapping and sequencing.

Seven of eight variants studied carried similar inversions between the *rrnD* operon and the *metE* gene. The mystifying case of strain AM2131 will be discussed below. The high level of *udp* gene expression in these strains (studied in detail by Kulakauskas et al. [17] for CM973) is due to read-through transcription from the promoters of the *rrnD* operon.

Physical mapping of chromosomal rearrangement in the

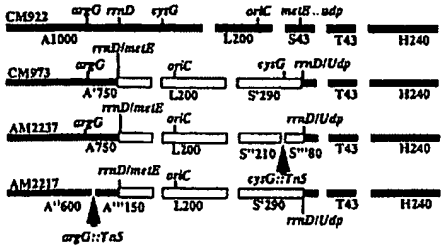

FIG. 3 Physical maps of the *arg-udp* region of the *E coli* strains derived from HfrH. The bars constituting the map represent *Not*I fragments. Their sizes and names (according to Riley and Krawiec [26]) are positioned under the bars. The locations of Tn*5*, which splits the *Not*I fragments, are shown by the arrows. The names of the strains are positioned at the beginning of each line. Empty bars represent inverted regions.

derivatives of AM2150 (W3110). The results of blot hybridizations of the *udp* and *rrn* probes with *Not*I fragments of chromosomal DNA isolated from strains AM2151 to AM2154 (derivatives of AM2150) are presented in Fig. 4, which reveals a picture more complicated than that with the derivatives of CM922. The precise linkage of the *rrn* and *metE-udp* regions in the chromosome of AM2151, as in the case of CM973, was confirmed by PCR mapping and sequencing (see below). To explain the blot hybridization data, we proposed a two-step scheme. The first step of the rearrangements is an inversion between the insertion *metE*::Tn*10* and the hybrid *rrnD/E*

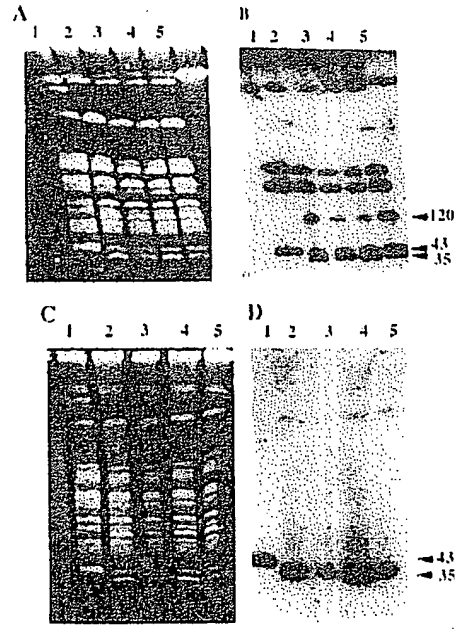

FIG. 4. Localization of the *udp* and *rrn* genes on the *Not*I restriction fragments of different *E. coli* strains. (A and C) PFGE separation; (B) blot hybridization with the *rrn* probe; (D) blot hybridization with the *udp* probe. Lanes: 1, AM2150; 2, AM2151; 3, AM2152; 4, AM2153; 5, AM2154

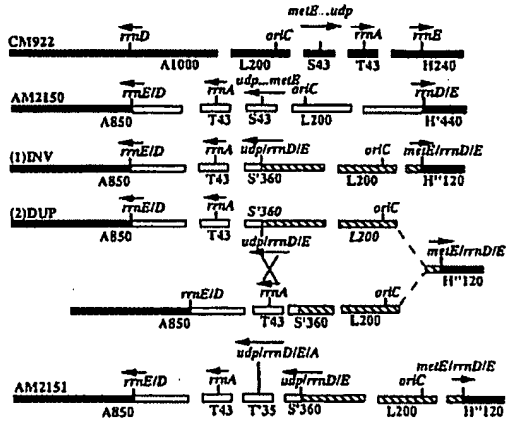

FIG. 5 Physical maps of the *arg-udp* region of the *E. coli* strains derived from W3110. The bars constituting the maps represent *Not*I fragments. Their sizes and names (according to Riley and Krawiec [26]) are positioned under the bars. The orientation of transcription of the important genes is shown by the arrows. The names of the strains are positioned at the beginning of each line. (1)INV and (2)DUP are the proposed intermediates. Empty bars represent a region originally inverted in the strain W3110 (AM2150). Striped bars represent the *rrn-udp* inversion.

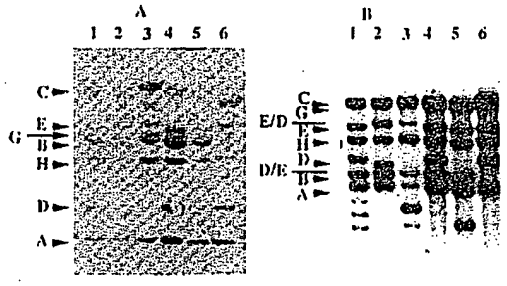

FIG. 6 Visualization of the restriction fragments harboring *rrn* operons in the rearranged strains of *E. coli*. (A) *Bam*HI digestion; (B) *Bam*HI-*Pst*I digestion. Lanes: 1, AB1157; 2, AM2150; 3, AM2151; 4, AM174; 5, CM973; 6, AM2131. Plasmid pKK3535 was used as the probe. Bands A, B, and C, etc., correspond to the operons *rrnA*, *rrnB*, *rrnC*, etc., respectively; bands D/E, etc., correspond to hybrid operons *rrnD/E*, etc.

operon (Fig. 5). The second step is a reciprocal crossover between *rrnA* on one sister chromatid and the hybrid *udp-rrnD/E* operon on the other. Finally, a region between *rrnA* and *udp-rrnD/E* was duplicated.

Thus, selection for a high level of *udp* gene expression either in AM2150 or in the CM922 background generated extended inversions in which different rRNA operons were involved. The direction of these inversions is dependent on the relative orientation of the *udp* gene and the particular *rrn* operon involved in the recombination.

Physical mapping of *INV*160 inversion in AM2131. To study the chromosomal structure of strain AM2131, another physical mapping technique was used. It has been shown elsewhere (3) that none of the seven *rrn* sequences in the *E. coli* contains either a *Bam*HI or a *Pst*I site. Therefore, if chromosomal DNA of *E. coli* is cleaved with these enzymes, seven fragments with distinct sizes can be detected by Southern hybridization with an *rrn* probe, each corresponding to a single *rrn* operon. We already knew that the *Not*I fragments carrying *rrn* operons in strain AM2131 are rearranged. If recombination events producing these rearrangements utilize *rrn* DNA, the method described by Boros et al. (3) can be used for mapping. Figure 6 presents a comparison of the *rrn* patterns of the mutant AM2131 and the parent strains. The data shown indicate that the parental strain AM174 has exactly the same pattern of *rrn* bands as the strain already mapped, strain AB1157. Strain AM2150 contains two hybrid bands: *rrnD/E* and *rrnE/D*. In contrast to the other mutants, strain AM2131 did not contain *rrnB* but contained a new band with a higher molecular weight.

The data presented in Fig. 6 and also the results of PFGE mapping allow us to suggest the following model of the chromosomal rearrangement in the AM2131 mutant (Fig. 7): first, an inversion between the *rrnG* and *rrnB* operons occurred, and, second, an inversion between the *metE* and hybrid *rrnG/B* operon fused the *udp* gene to the *rrnG/B* promoter.

We do not yet have direct sequencing data on the structure of the last rearrangement; however, changes in the mobilities of the restriction *Bam*HI and *Bam*HI-*Pst*I fragments mostly consisting of *rrn* DNA make the proposed scheme very probable.

In summary, 8 of 14 variants selected from strain CM922 and all four variants of AM2150 have the *udp* gene fused to the *rrn* promoter. In five strains, the initial inversion event induced further rearrangements.

PCR mapping and sequencing of rearranged mutants. Precise mapping of the rearrangement breakpoints was performed by PCR amplification of the DNA fragments connected by the proposed inversions. In these amplification experiments, represented in Fig. 8, sequences near each end of the IS*10* (IS_p and IS_d) were used as primers in combination with a set of primers specific to the *rrn* operons and the *metE* gene. Chromosomal DNA specimens from CM973 and AM2151 (chosen to represent each group of rearranged strains) were used as templates. Among all pairs of primers used, primers rrn7, rrn8, and rrn10 combined with IS_d, along with $metE_p$ and $metE_d$ combined with IS_p, produced visible amplification fragments (Fig. 8). The existence of the DNA fragments linking IS*10* DNA and *rrn* DNA confirms the proposed scheme of recombination. To establish the exact endpoints of the *rrnD*-IS*10* fusion, both strands of the PCR-generated fragments were sequenced. In the chromosome of CM973, nucleotide 3699 of

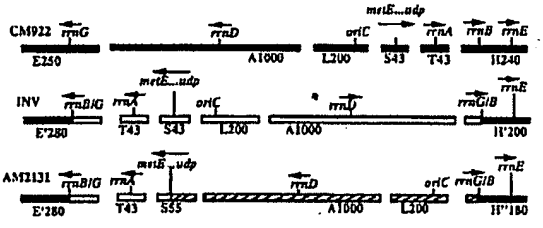

FIG. 7. Possible scheme of chromosomal rearrangements in strain AM2131. The bars constituting the maps represent *Not*I fragments. Their sizes and names (according to Riley and Krawiec [26]) are positioned under the bars. The orientation of transcription of the important genes is shown by arrows. The names of the strains are positioned at the beginning of each line. INV is the proposed intermediate. Empty bars represent a region (*rrnB-rrnG*) initially inverted in the proposed intermediate. Striped bars represent the second *rrn-udp* inversion.

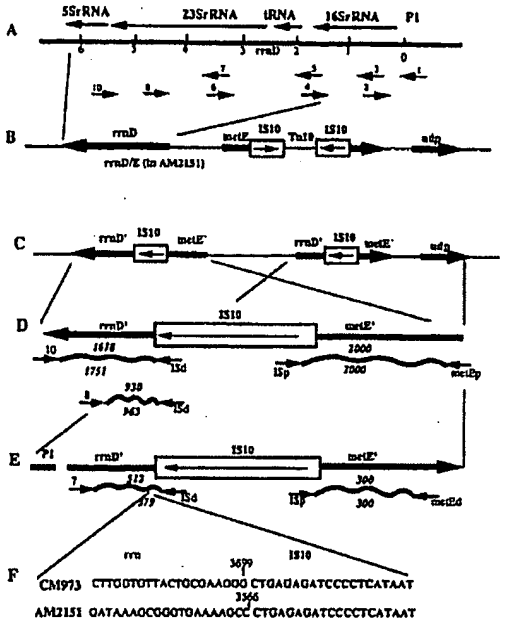

FIG. 8. Scheme of PCR mapping and sequencing of the rearrangement endpoints. (A) *rrnD* operon (arrows with numbers represent primers used for PCR mapping); (B) region of the *E. coli* chromosome studied (bars with arrows represent orientation of transcription of the genes described in the text; arrows inside boxes represent the orientation of IS*10*); (C) chromosome after rearrangements; (D and E) enlarged regions around rearrangement endpoints (the wavy lines represent PCR products; the numbers above them represent their sizes for CM973; the numbers below them represent their sizes for AM2151; arrows represent primers used for PCR mapping; (F) DNA sequences of the *rrn*-IS*10* rearrangement endpoints for CM973 and AM2151.

rrnD was merged with the very beginning of the IS*10* that is connected with *metE-udp*; in AM2151, the corresponding nucleotide is 3566 (Fig. 8). Comparison of the rearrangement endpoints in two strains studied did not reveal any consensus in DNA sequences; however, both of them are separated by only 133 nucleotides. Sequencing more *rrn*-IS*10* junctions will provide material for more-detailed conclusions

DISCUSSION

Selection for overexpression of the *udp* gene (with Tn*10* integrated upstream) led to its fusion to the *rrn* promoter in more than half of the variants analyzed. Our earlier experiments (1) demonstrated that a duplication of the *udp* gene is sufficient to provide the expression enhancement selectable by the proposed scheme. Therefore, the system does not need great overexpression, and it is not clear what is so unique in the *rrn* promoter to make it so attractive for the *udp* gene (i.e., to fuse with it preferably). In a related selection system, one spontaneous inversion was found in 1 of 10,000 cases that were analyzed (28). In another work, pieces of homologous DNA were introduced in the proposed endpoints (19), and directed inversions then contributed 65 to 90% of all selected variants. This is very close to the results that we obtained (about 70%), although the mechanisms of the inversions are probably different. Direct sequence comparison of Tn*10* and the *rrn* operon does not provide enough homology to explain the rearrangements selected in the study by the mechanism of homologous recombination (5).

Five of 12 inversions studied were accompanied by duplications of the chromosomal segment flanked by *rrn* operons. Similar rearrangements were demonstrated for the *rrnB-rrnE* region of the *E. coli* chromosome, where the probability of recombination varied from 10^{-3} for the deletion to 5% for the duplications (in the UV-irradiated cells [for a review, see reference 13]).

One of the key elements of the system is Tn*10*, which by itself can generate inversions by intramolecular transposition involving the inside ends of its IS*10* elements (15, 30). The results of the sequencing of the PCR products linking IS*10* of Tn*10* and *rrn* DNA are consistent with the scheme of the inversion formation, emphasizing the role of the IS*10*. More sequencing data generated from other endpoints may reveal the target specificity of this process. However, another possible explanation of the clustering of the endpoints is that it is not due to specific target sequences within the *rrn* DNA but is rather due to the conserved three-dimensional structure of the bacterial chromosome. This will make some rearrangements more probable than others, limiting selected inversions to certain preferable areas (27). In this case, more sequencing may not reveal target similarities in the inversion endpoints selected in the work.

Another possible extension of this work is the study of the conservation of the gene order in the bacterial chromosome. The only direct comparison of whole chromosomal maps, which has been done for three different *Bacillus* species (6), revealed numerous gene rearrangements that contradict the observations accumulated by studies of members of the family *Enterobacteriaciae*. The low resolution of the mapping in this work (6) does not allow final conclusions, but it conflicts with well-established views of the bacterial chromosome as a strongly conserved structure. Similar results were derived from analysis of the high-resolution physical and genetic maps of *Rhodobacter capsulatus* SB1003 (8), which is now being expanded to three more strains. In comparisons of these maps, a set of long-range rearrangements has been revealed. Some of these have endpoints in or near the *rrn* operons.

Inversion strains generated in the present work have reduced growth rates (data not shown). Selection for faster-growing variants can be used for the modeling of genome-stabilizing rearrangements, as was done by Hill and Gray (11). One can expect to select reinversions in the process of genome stabilization; however, the mechanisms should be different from the ones described, because the initial triggering event is different.

ACKNOWLEDGMENTS

We thank R. Haselkorn and V. Vonstein for their critical reading of the manuscript. We are grateful to D. Berg, J. Roth, and M. Schmid for helpful discussions; C. Berg, J. Cole, S. Brown, and S. Mindlin for providing strains; V. P. Veiko for providing oligonucleotides, and G. Sergeeva for valuable technical assistance.

This work was supported by Fogarty Foundation grant TW00039, by PHS grant HG00563, and by the Russian State Program Frontiers in Genetics grant 553/92.

REFERENCES

1. **Alkhimova, R. A., V. V. Sukhodolets, and A. S. Mironov.** 1985. Enhancement of expression of Escherichia coli uridine phosphorylase gene as a result of duplication. Genetika 21:756–762.
2. **Bachmann, B.** 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12. p. 1190–1219. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.) *Escherichia coli* and *Salmonella*

typhimurium: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
3. **Boros, I., A. Kiss, and P. Venetianer.** 1979. Physical map of the seven ribosomal RNA genes of Escherichia coli. Nucleic Acids Res. 6:1817–1830.
4. **Brikun, I. A., A. S. Mironov, R. V. Masiliunaite, and V. V. Sukhodolets.** 1990. Cloning of Escherichia coli uridine phosphorylase gene: localization of structural and regulatory regions in the cloned fragment and identification of the protein product. Mol. Genet. Mikrobiol. Virusol. 6:7–11.
5. **Brosius, J., T. J. Dull, D. D. Sleeter, and H. F. Noller.** 1981. Gene organization and primary structure of a ribosomal RNA operon from Escherichia coli. J. Mol. Biol. 148:107–127.
6. **Carlson, C. R., A. Gronstad, and A. B. Kolsto.** 1992. Physical maps of the genomes of three *Bacillus cereus* strains. J. Bacteriol. 174:3750–3756.
7. **Daniel, D. L.** 1990. Constructing encyclopedias of genomes, p. 43–51. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
8. **Fonstein, M., and R. Haselkorn.** 1993. Chromosomal structure of *Rhodobacter capsulatus* strain SB 1003: cosmid encyclopedia and high resolution physical and genetic map. Proc. Natl. Acad. Sci. USA 90:2522–2526.
9. **Fonstein, M., S. Zheng, and R. Haselkorn.** 1992. Physical map of the genome of *Rhodobacter capsulatus* SB 1003. J. Bacteriol. 174:4070–4077.
10. **Hartman, P. E., Z. Hartman, and R. C. Stahl.** 1971. Classification and mapping of spontaneous and induced mutations in the histidine operon of Salmonella. Adv. Genet. 16:1–34.
11. **Hill, C. W., and J. A. Gray.** 1988. Effects of chromosomal inversion on cell fitness in Escherichia coli K-12. Genetics 119:771–778.
12. **Hill, C. W., and B. W. Harnish.** 1981. Inversions between ribosomal RNA genes of Escherichia coli. Proc. Natl. Acad. Sci. USA 78:7069–7072.
13. **Hill, C. W., S. Harvey, and J. A. Gray.** 1990. Recombination between rRNA genes in *Escherichia coli* and *Salmonella typhimurium*, p. 335–340. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
14. **Kingston, R. E., R. R. Gutell, A. R. Taylor, and M. J. Chamberlin.** 1981. Transcriptional mapping of plasmid pKK3535. Quantitation of the effect of guanosine tetraphosphate on binding to the rrnB promoters and a lambda promoter with sequence homologies in the CII binding region. J. Mol. Biol. 146:433–449.
15. **Kleckner, N., K. Reichardt, and D. Botstein.** 1979. Inversions and deletions of the Salmonella chromosome generated by the translocatable tetracycline resistance element Tn10. J. Mol. Biol. 127:89–115.
16. **Konrad, E. B.** 1977. Method for the isolation of *Escherichia coli* mutants with enhanced recombination between chromosomal duplications. J. Bacteriol. 130:167–172.
17. **Kulakauskas, S. T., V. V. Sukhodolets, and A. S. Mironov.** 1985. Chromosomal inversion accompanied by an enhancement of uridine phosphorylase gene expression in Escherichia coli K-12. Genetika 21:375–383.
18. **Low, K. B.** 1987. Hfr strains of *E. coli*, p. 1071–1109. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
19. **Mahan, M. J., A. M. Segall, and J. R. Roth.** 1990. Recombination events that rearrange the chromosome; barriers to inversion, p. 341–350. *In* K. Drlica and M. Riley (ed.), The bacterial chromosome. American Society for Microbiology, Washington, D.C.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Mironov, A. S., and V. V. Sukhodolets.** 1979. Promoter-like mutants with increased expression of the *Escherichia coli* uridine phosphorylase structural gene. J. Bacteriol. 137:802–810.
23. **Murray, V.** 1989. Improved double-stranded DNA sequencing using the linear polymerase chain reaction. Nucleic Acids Res. 17:2675.
24. **Rebollo, J. E., V. Francois, and J. M. Louarn.** 1988. Detection and possible role of two large nondivisible zones on the Escherichia coli chromosome. Proc. Natl. Acad. Sci. USA 85:9391–9395.
25. **Riley, M., and A. Anilionis.** 1978. Evolution of the bacterial genome. Annu. Rev. Microbiol. 32:519–560.
26. **Riley, M., and S. Krawiec.** 1987. Genome organization, p. 967–981. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
27. **Schmid, M. B., P. J. Krug, A. Z. Gileski, R. J. Code, and A. Torjussen.** 1993. Genetic method reveals nucleoid structural information?, p. 12. Second International *E. coli* Genome Meeting, Madison, Wis.
28. **Schmid, M. B., and J. R. Roth.** 1983. Genetic methods for analysis and manipulation of inversion mutations in bacteria. Genetics 105:517–537.
29. **Segall, A., M. J. Mahan, and J. R. Roth.** 1988. Rearrangement of the bacterial chromosome: forbidden inversions. Science 241: 1314–1318.
30. **Shen, M. M., E. A. Raleigh, and N. Kleckner.** 1987. Physical analysis of Tn10- and IS10-promoted transpositions and rearrangements. Genetics 116:359–369.
31. **Smith, C. L., J. G. Econome, A. Schutt, S. Klco, and C. R. Cantor.** 1987. A physical map of the Escherichia coli K12 genome. Science 236:1448–1453.
32. **Smith, C. L., and R. D. Kolodner.** 1988. Mapping of Escherichia coli chromosomal Tn5 and F insertions by pulsed field gel electrophoresis. Genetics 119:227–236.
33. **Young, K. K., G. J. Edlin, and G. G. Wilson.** 1982. Genetic analysis of bacteriophage T4 transducing bacteriophages. J. Virol. 41:345–347.

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☑ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.